

04 20



#2

OIKE

RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/10/071,976

TIME: 16:16:52

Input Set : N:\Crf3\RULE60\10071976.txt

Output Set: N:\CRF3\02252002\J071976.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Tsien et al., Roger Y. et al.

7 (ii) TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT
8 PROTEINS

10 (iii) NUMBER OF SEQUENCES: 4

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Fish & Richardson P.C.

14 (B) STREET: 4225 Executive Square, Suite 1400

15 (C) CITY: La Jolla

16 (D) STATE: CA

17 (E) COUNTRY: USA

18 (F) ZIP: 92037

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

26 (vi) CURRENT APPLICATION DATA:

C--> 27 (A) APPLICATION NUMBER: US/10/071,976

C--> 28 (B) FILING DATE: 05-Feb-2002

33 (C) CLASSIFICATION:

34 (vii) PRIOR APPLICATION DATA:

31 (A) APPLICATION NUMBER: 09/465,142

32 (B) FILING DATE: 1999-DEC-16

35 (A) APPLICATION NUMBER: 08/974,737

36 (B) FILING DATE:

38 (viii) ATTORNEY/AGENT INFORMATION:

39 (A) NAME: Haile, Lisa A.

40 (B) REGISTRATION NUMBER: 38,347

41 (C) REFERENCE/DOCKET NUMBER: 07257/056001

43 (ix) TELECOMMUNICATION INFORMATION:

44 (A) TELEPHONE: 619/678-5070

45 (B) TELEFAX: 619/678-5099

48 (2) INFORMATION FOR SEQ ID NO: 1:

50 (i) SEQUENCE CHARACTERISTICS:

51 (A) LENGTH: 716 base pairs

52 (B) TYPE: nucleic acid

53 (C) STRANDEDNESS: single

54 (D) TOPOLOGY: linear

58 (ix) FEATURE:

59 (A) NAME/KEY: CDS

60 (B) LOCATION: 1..714

ENTERED

RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/10/071,976

TIME: 16:16:52

Input Set : N:\Crif3\RULE60\10071976.txt

Output Set: N:\CRF3\02252002\J071976.raw

```

63      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
65 ATG AGT AAA GGA GAA GAA CTT TTC ACT GCA GTT GTC CCA ATT CTT GTT      48
66 Met Ser Lys Gly Glu Glu Leu Phe Thr Ala Val Val Pro Ile Leu Val
67   1           5           10           15
69 GAA TTA GAT GGT GAT GTT AAT GGG CAC AAA TTT TCT GTC AGT GGA GAG      96
70 Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly Glu
71           20           25           30
73 GGT GAA GGT GAT GTA ACA TAC GGA AAA CTT ACC CTT AAA TTT ATT TGC      144
74 Gly Glu Gly Asp Val Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile Cys
75           35           40           45
77 ACT ACT GGA AAA CTA CCT GTT CCA TGG CCA ACA CTT GTC ACT ACT TTC      192
78 Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr Phe
79   50           55           60
81 TCT TAT GGT GTT CAA TGC TTT TCA AGA TAC CCA GAT CAT ATG AAA CGG      240
82 Ser Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His Met Lys Arg
83 65           70           75           80
87 CAT GAC TTT TTC AAG AGT GCC ATG CCC GAA GGT TAT GTA CAG CAA AGA      288
88 His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Gln Arg
89           85           90           95
91 ACT ATA TTT TTC AAA GAT GAC GGG AAC TAC AAG ACA CGT GCT GAA GTC      336
92 Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val
93           100          105          110
95 AAG TTT GAA GGT GAT ACC CTT GTT AAT AGA ATC GAG TTA AAA GGT ATT      384
96 Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile
97           115          120          125
99 GAT TTT AAA GAA GAT GGA AAC ATT CTT GGA CAT AAA TTG GAA TAC AAC      432
100 Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn
101           130          135          140
103 TAT AAC TCA CAC AAT GTA TAC ATC ATG GCA GAC AAA CAA AAG AAT GGA      480
104 Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly
105 145          150          155          160
107 ATC AAA GTT AAC TTC AAA ATT AGA CAC AAC ATT GAA GAT GGA AGC GTT      528
108 Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser Val
109           165          170          175
111 CAA CTA GCA GAC TAT TAT CAA CAA AAT ACT CCA ATT CTC GAT GGC CCT      576
112 Gln Leu Ala Asp Tyr Tyr Gln Gln Asn Thr Pro Ile Leu Asp Gly Pro
113           180          185          190
115 GTC CTT TTA CCA GAC AAC CAT TAC CTG TCC ACA CAA TCT GCC CTT TCG      624
116 Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser Ala Leu Ser
117           195          200          205
119 AAA GAT CCC AAC GAA AAG AGA GAC CAC ATG GTC CTT CTT GAG TTT GTA      672
120 Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe Val
121           210          215          220
123 ACA GCT GCT GGG ATT ACA CAT GGC ATG GAT GAA CTA TAC AAA      714
124 Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
125 225          230          235
127 TA      716
130 (2) INFORMATION FOR SEQ ID NO: 2:
132      (i) SEQUENCE CHARACTERISTICS:

```

RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/10/071,976

TIME: 16:16:52

Input Set : N:\Crf3\RULE60\10071976.txt

Output Set: N:\CRF3\02252002\J071976.raw

133 (A) LENGTH: 238 amino acids
 134 (B) TYPE: amino acid
 135 (D) TOPOLOGY: linear
 137 (ii) MOLECULE TYPE: protein
 139 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

141	Met	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Ala	Val	Val	Pro	Ile	Leu	Val
142	1				5					10					15	
144	Glu	Leu	Asp	Gly	Asp	Val	Asn	Gly	His	Lys	Phe	Ser	Val	Ser	Gly	Glu
145				20				25					30			
147	Gly	Glu	Gly	Asp	Val	Thr	Tyr	Gly	Lys	Leu	Thr	Leu	Lys	Phe	Ile	Cys
148				35				40					45			
150	Thr	Thr	Gly	Lys	Leu	Pro	Val	Pro	Trp	Pro	Thr	Leu	Val	Thr	Thr	Phe
151		50					55					60				
153	Ser	Tyr	Gly	Val	Gln	Cys	Phe	Ser	Arg	Tyr	Pro	Asp	His	Met	Lys	Arg
154	65					70				75					80	
156	His	Asp	Phe	Phe	Lys	Ser	Ala	Met	Pro	Glu	Gly	Tyr	Val	Gln	Gln	Arg
157					85					90					95	
159	Thr	Ile	Phe	Phe	Lys	Asp	Asp	Gly	Asn	Tyr	Lys	Thr	Arg	Ala	Glu	Val
160				100					105					110		
162	Lys	Phe	Glu	Gly	Asp	Thr	Leu	Val	Asn	Arg	Ile	Glu	Leu	Lys	Gly	Ile
163			115						120					125		
165	Asp	Phe	Lys	Glu	Asp	Gly	Asn	Ile	Leu	Gly	His	Lys	Leu	Glu	Tyr	Asn
166		130					135					140				
168	Tyr	Asn	Ser	His	Asn	Val	Tyr	Ile	Met	Ala	Asp	Lys	Gln	Lys	Asn	Gly
169	145					150					155					160
171	Ile	Lys	Val	Asn	Phe	Lys	Ile	Arg	His	Asn	Ile	Glu	Asp	Gly	Ser	Val
172				165					170						175	
174	Gln	Leu	Ala	Asp	Tyr	Tyr	Gln	Gln	Asn	Thr	Pro	Ile	Leu	Asp	Gly	Pro
175			180						185					190		
177	Val	Leu	Leu	Pro	Asp	Asn	His	Tyr	Leu	Ser	Thr	Gln	Ser	Ala	Leu	Ser
178			195						200					205		
180	Lys	Asp	Pro	Asn	Glu	Lys	Arg	Asp	His	Met	Val	Leu	Leu	Glu	Phe	Val
181		210					215					220				
183	Thr	Ala	Ala	Gly	Ile	Thr	His	Gly	Met	Asp	Glu	Leu	Tyr	Lys		
184	225					230					235					

186 (2) INFORMATION FOR SEQ ID NO: 3:

188 (i) SEQUENCE CHARACTERISTICS:

189 (A) LENGTH: 720 base pairs

190 (B) TYPE: nucleic acid

191 (C) STRANDEDNESS: single

192 (D) TOPOLOGY: linear

196 (ix) FEATURE:

197 (A) NAME/KEY: CDS

198 (B) LOCATION: 1..720

201 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

203	ATG	GTG	AGC	AAG	GGC	GAG	GAG	CTG	TTC	ACC	GGG	GTG	GTG	CCC	ATC	CTG	48
204	Met	Val	Ser	Lys	Gly	Glu	Glu	Leu	Phe	Thr	Gly	Val	Val	Pro	Ile	Leu	
W--> 205	240					245						250					
207	GTC	GAG	CTG	GAC	GGC	GAC	GTA	AAC	GGC	CAC	AAG	TTC	AGC	GTG	TCC	GGC	96

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/071,976

DATE: 02/25/2002

TIME: 16:16:52

Input Set : N:\Crf3\RULE60\10071976.txt

Output Set: N:\CRF3\02252002\J071976.raw

```

208 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
W--> 209 255                260                265                270
211 GAG GGC GAG GGC GAT GCC ACC TAC GGC AAG CTG ACC CTG AAG TTC ATC      144
212 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
W--> 213                275                280                285
215 TGC ACC ACC GGC AAG CTG CCC GTG CCC TGG CCC ACC CTC GTG ACC ACC      192
216 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
W--> 217                290                295                300
219 TTC GGC TAC GGC GTG CAG TGC TTC GCC CGC TAC CCC GAC CAC ATG AAG      240
220 Phe Gly Tyr Gly Val Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys
W--> 221                305                310                315
223 CAG CAG GAC TTC TTC AAG TCC GCC ATG CCC GAA GGC TAC GTC CAG GAG      288
224 Gln Gln Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
W--> 225                320                325                330
227 CGC ACC ATC TTC TTC AAG GAC GAC GGC AAC TAC AAG ACC CGC GCC GAG      336
228 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
W--> 229 335                340                345                350
231 GTG AAG TTC GAG GGC GAC ACC CTG GTG AAC CGC ATC GAG CTG AAG GGC      384
232 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
W--> 233                355                360                365
235 ATC GAC TTC AAG GAC GAC GGC AAC ATC CTG GGG CAC AAG CTG GAG TAC      432
236 Ile Asp Phe Lys Asp Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
W--> 237                370                375                380
239 AAC TAC AAC AGC CAC AAC GTC TAT ATC ATG GCC GAC AAG CAG AAG AAC      480
240 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
W--> 241                385                390                395
243 GGC ATC AAG GTG AAC TTC AAG ATC CGC CAC AAC ATC GAG GAC GGC AGC      528
244 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
W--> 245                400                405                410
249 GTG CAG CCC GCC GAC CAC TAC CAG CAG AAC ACC CCC ATC GGC GAC GGC      576
250 Val Gln Pro Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
W--> 251 415                420                425                430
253 CCC GTG CTG CTG CCC GAC AAC CAC TAC CTG AGC TAC CAG TCC GCC CTG      624
254 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
W--> 255                435                440                445
257 AGC AAA GAC CCC AAC GAG AAG CGC GAT CAC ATG GTC CTG CTG GAG TTC      672
258 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
W--> 259                450                455                460
261 GTG ACC GCC GCC GGG ATC ACT CAC GGC ATG GAC GAG CTG TAC AAG TAA      720
262 Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys *
W--> 263                465                470                475
266 (2) INFORMATION FOR SEQ ID NO: 4:
268     (i) SEQUENCE CHARACTERISTICS:
269         (A) LENGTH: 239 amino acids
270         (B) TYPE: amino acid
271         (D) TOPOLOGY: linear
273     (ii) MOLECULE TYPE: protein
275     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
277 Met Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro Ile Leu

```

RAW SEQUENCE LISTING

DATE: 02/25/2002

PATENT APPLICATION: US/10/071,976

TIME: 16:16:52

Input Set : N:\Crf3\RULE60\10071976.txt

Output Set: N:\CRF3\02252002\J071976.raw

```

278      1              5              10              15
280 Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly
281              20              25              30
283 Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile
284              35              40              45
286 Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr
287              50              55              60
289 Phe Gly Tyr Gly Val Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys
290      65              70              75              80
292 Gln Gln Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu
293              85              90              95
295 Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu
296              100              105              110
298 Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly
299              115              120              125
301 Ile Asp Phe Lys Asp Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr
302              130              135              140
304 Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn
305      145              150              155              160
307 Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser
308              165              170              175
310 Val Gln Pro Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp Gly
311              180              185              190
313 Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu
314              195              200              205
316 Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe
317              210              215              220
319 Val Thr Ala Ala Gly Ile Thr His Gly Met Asp Glu Leu Tyr Lys
320      225              230              235

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/071,976

DATE: 02/25/2002

TIME: 16:16:53

Input Set : N:\Crif3\RULE60\10071976.txt

Output Set: N:\CRF3\02252002\J071976.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:205 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:209 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:213 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:217 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:221 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:225 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:229 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:233 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:237 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:241 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:245 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:251 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:255 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:259 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3
L:263 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:3